GenCore version 4.5


OM nucleic - nucleic search, using sw model

Run on: February 26, 2002, 11:48:04 ; Search time 205.96 Seconds
(without alignments)
4645.441 Million cell updates/sec

Title: US-09-602-833A-1
Perfect score: 1116
Sequence: 1 atgggacataaagtggttgt.....cttttagccttcaactttga 1116

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 930621 seqs, 428662619 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_1101:*
1: /SIDS2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| | Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|---|
| | 1 | 1116 | 100.0 | 1116 | 22 | AAF24902 | Nucleotide sequenc |
| | 2 | 681 | 61.0 | 681 | 22 | AAF24903 | Nucleotide sequenc |
| | 3 | 129 | 11.6 | 2056 | 22 | AAH17218 | Human cDNA sequenc |
| | 4 | 23 | 2.1 | 9789 | 17 | AAT41852 | cDNA encoding Plas |
| c | 5 | 20 | 1.8 | 1421 | 21 | AAC46422 | Arabidopsis thalia |
| c | 6 | 20 | 1.8 | 1422 | 21 | AAC36467 | Arabidopsis thalia |
| | 7 | 19 | 1.7 | 327 | 21 | AAC00689 | Human secreted pro |
| | 8 | 19 | 1.7 | 413 | 21 | AAC06478 | Human secreted pro |
| | 9 | 19 | 1.7 | 1185 | 16 | AAT03478 | Transcription fact |
| | 10 | 19 | 1.7 | 1816 | 21 | AAC78156 | Human cancer assoc |
| | 11 | 19 | 1.7 | 3072 | 21 | AAA75580 | DNA encoding a mou |
| | 12 | 18 | 1.6 | 811 | 19 | AAX13947 | H. pylori GHPO 127 |
| | 13 | 18 | 1.6 | 1981 | 18 | AAT66241 | Romaine lettuce vi |
| c | 14 | 18 | 1.6 | 2324 | 22 | AAD10125 | Mouse serotransfer |
| c | 15 | 18 | 1.6 | 3997 | 19 | AAV26082 | Tomato pest resist |
| c | 16 | 18 | 1.6 | 4118 | 17 | AAT44520 | NTHI hxuC + hxuB g |
| c | 17 | 18 | 1.6 | 4566 | 19 | AAV41550 | Nucleotide sequenc |
| c | 18 | 18 | 1.6 | 4566 | 20 | AAX01089 | Human G-protein co |
| c | 19 | 18 | 1.6 | 4568 | 18 | AAT44039 | Human G-protein re |
| c | 20 | 18 | 1.6 | 4568 | 20 | AAX01094 | Human G-protein co |
| | 21 | 18 | 1.6 | 6253 | 20 | AAX13097 | Enterococcus faeca |
| c | 22 | 18 | 1.6 | 11648 | 22 | AAD08065 | Human extracellula |
| c | 23 | 18 | 1.6 | 51952 | 19 | AAV26084 | Tomato pest resist |
| c | 24 | 18 | 1.6 | 68940 | 20 | AAX57351 | Human chromosome 6 |
| c | 25 | 18 | 1.6 | 1038602 | 20 | AAZ01425 | Complete genome se |
| c | 26 | 17 | 1.5 | 156 | 22 | AAI19482 | Probe #9415 for ge |
| c | 27 | 17 | 1.5 | 156 | 22 | AAI20472 | Probe #10405 for g |
| c | 28 | 17 | 1.5 | 156 | 22 | AAI44677 | Probe #13363 used |
| c | 29 | 17 | 1.5 | 156 | 22 | AAI45679 | Probe #14365 used |
| c | 30 | 17 | 1.5 | 156 | 22 | AAI05210 | Probe #5201 used t |
| c | 31 | 17 | 1.5 | 156 | 22 | AAI06170 | Probe #6161 used t |
| | 32 | 17 | 1.5 | 166 | 21 | AAA41702 | Human secreted exp |
| | 33 | 17 | 1.5 | 179 | 21 | AAC17644 | Human secreted pro |
| | 34 | 17 | 1.5 | 243 | 21 | AAC17113 | Human secreted pro |
| | 35 | 17 | 1.5 | 291 | 21 | AAF08533 | Fusarium venenatum |
| | 36 | 17 | 1.5 | 296 | 18 | AAV78820 | Staphylococcus aur |
| c | 37 | 17 | 1.5 | 390 | 21 | AAA67203 | Pinus radiata alph |
| c | 38 | 17 | 1.5 | 418 | 21 | AAA67201 | Pinus radiata alph |
| c | 39 | 17 | 1.5 | 441 | 21 | AAA67198 | Pinus radiata alph |
| | 40 | 17 | 1.5 | 443 | 22 | AAH35681 | Human colon cancer |
| | 41 | 17 | 1.5 | 460 | 22 | AAI15848 | Probe #5781 for ge |
| | 42 | 17 | 1.5 | 460 | 22 | AAI37733 | Probe #6419 used t |
| | 43 | 17 | 1.5 | 476 | 22 | AAI16028 | Probe #5961 for ge |
| | 44 | 17 | 1.5 | 476 | 22 | AAI38141 | Probe #6827 used t |
| c | 45 | 17 | 1.5 | 479 | 22 | AAI10199 | Probe #132 for gen |

ALIGNMENTS

RESULT 1
AAF24902
ID AAF24902 standard; cDNA; 1116 BP.
XX
AC AAF24902;
XX
DT 20-APR-2001 (first entry)
XX
DE Nucleotide sequence of a human SGT4-1 polypeptide.
XX
KW Human; SGT4; signal transduction; guanosine triphosphate binding protein;
KW GTP binding protein; cancer; immune response; nutritional source;
KW animal feed; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1116
FT /*tag= a
FT /product= "SGT4"
XX
PN WO200078959-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-US17248.
XX
PR 23-JUN-1999; 99US-0140627.
XX
PA (LEXI-) LEXICON GENETICS INC.
XX
PI Turner AC, Zambrowicz B, Nehls M, Friedrich GA, Sands AT;
XX
DR WPI; 2001-032329/04.

```
DR   P-PSDB; AAB31563.
XX
PT   New SGT4 genes and proteins, useful for diagnosing and treating
PT   disorders involving inappropriate regulation of a signal transduction
PT   mechanism e.g. cancer -
XX
PS   Claim 1; Fig 1; 82pp; English.
XX
CC   The present sequence encodes a human SGT4 polypeptide. SGT4 polypeptides
CC   are involved in signal transduction pathways regulated by guanosine
CC   triphosphate (GTP) binding proteins). SGT4 polynucleotides and
CC   polypeptides are for diagnosing and treating conditions related to a
CC   signal transduction mechanism involving SGT4 such as cancer. In
CC   addition, it can be used to detect the expression of SGT4 as markers of
CC   specific cells and tissues such as neuronal tissue, heart, liver,
CC   pancreas and adrenal gland. They are also useful for the construction of
CC   transgenic and knockout animals for studying SGT4 function in vivo and
CC   for the screening of SGT4 (ant)agonists in an animal model. Other more
CC   general uses include: as molecular weight markers on Southern gels; as
CC   chromosome markers or tags; as probes; for selecting and making
CC   oligomers for attachment to a gene chip; to raise anti-protein or
CC   anti-DNA antibodies or to elicit immune response. They are also
CC   also be used as nutritional sources or supplements such as in animal
CC   feed.
XX
SQ   Sequence 1116 BP; 343 A; 224 C; 265 G; 284 T; 0 other;


  Query Match           100.0%;  Score 1116;  DB 22;  Length 1116;
  Best Local Similarity  100.0%;  Pred. No. 0;
  Matches 1116;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 atgggacataaagtggttgtcttcgacatttctgtcatcagagccttgtgggaaactcgt 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgggacataaagtggttgtcttcgacatttctgtcatcagagccttgtgggaaactcgt 60

Qy     61 gtcaagaagcacaaagcttggcagaagaaggaggtggaaaggcttgagaagagcgccttg 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 gtcaagaagcacaaagcttggcagaagaaggaggtggaaaggcttgagaagagcgccttg 120

Qy    121 gagaagataaaggaggagtggaactttgtggccgaatgcaggaggaagggcatcccccag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 gagaagataaaggaggagtggaactttgtggccgaatgcaggaggaagggcatcccccag 180

Qy    181 gctgtatactgcaagaatggcttcatagacaccagcgtgcggcttctggacaagattgaa 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 gctgtatactgcaagaatggcttcatagacaccagcgtgcggcttctggacaagattgaa 240

Qy    241 aggaacactctcacaaggcagagttcacttcccaaggacagaggcaaacggagcagtgcg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 aggaacactctcacaaggcagagttcacttcccaaggacagaggcaaacggagcagtgcg 300

Qy    301 tttgtgtttgaactttctggggagcactggacggagctcccagattcattgaaggagcag 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 tttgtgtttgaactttctggggagcactggacggagctcccagattcattgaaggagcag 360

Qy    361 acacacctgagagaatggtacataagcaataccttgattcaaatcattcctacatatatt 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 acacacctgagagaatggtacataagcaataccttgattcaaatcattcctacatatatt 420

Qy    421 cagttatttcaagcgatgagaattctggatctgccaaaaaaccaaatctcacatcttcca 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cagttatttcaagcgatgagaattctggatctgccaaaaaaccaaatctcacatcttcca 480

Qy    481 gcagaaatcggttgtttgaagaacctgaaagaactcaatgtgggtttcaactatctgaag 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 gcagaaatcggttgtttgaagaacctgaaagaactcaatgtgggtttcaactatctgaag 540

Qy    541 agcattcctccagaattgggagattgtgaaaatctagagagactggattgttctggaaat 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 agcattcctccagaattgggagattgtgaaaatctagagagactggattgttctggaaat 600

Qy    601 ctagaattaatggagctgccctttgaattaagtaatttgaagcaagttacatttgtagat 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 ctagaattaatggagctgccctttgaattaagtaatttgaagcaagttacatttgtagat 660

Qy    661 atctcagcaaacaagttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcag 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 atctcagcaaacaagttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcag 720

Qy    721 tggttggatatcagcagcaataacctgaccgacctgccgcaagatatagacaggctagag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 tggttggatatcagcagcaataacctgaccgacctgccgcaagatatagacaggctagag 780

Qy    781 gagctgcagagctttctcttgtataaaaacaagttgacctaccttccctattccatgctg 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 gagctgcagagctttctcttgtataaaaacaagttgacctaccttccctattccatgctg 840

Qy    841 aacctgaagaagctcactctgttagtcgtcagtggggaccatttggtggagctcccaact 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 aacctgaagaagctcactctgttagtcgtcagtggggaccatttggtggagctcccaact 900

Qy    901 gccctttgtgactcatccacacctttaaaatttgtaagccttatggacaatcctattgat 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 gccctttgtgactcatccacacctttaaaatttgtaagccttatggacaatcctattgat 960

Qy    961 aatgcccaatgtgaagatggcaatgaaataatggaaagtgaacgggatcgccaacatttt 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 aatgcccaatgtgaagatggcaatgaaataatggaaagtgaacgggatcgccaacatttt 1020

Qy   1021 gataaagaagttatgaaagcctatattgaagaccttaaagaaagagaatctgttcccagc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 gataaagaagttatgaaagcctatattgaagaccttaaagaaagagaatctgttcccagc 1080

Qy   1081 tataccaccaaagtgtcttttagccttcaactttga 1116
          ||||||||||||||||||||||||||||||||||||
Db   1081 tataccaccaaagtgtcttttagccttcaactttga 1116


RESULT   2
AAF24903
ID   AAF24903 standard; cDNA; 681 BP.
XX
AC   AAF24903;
XX
DT   20-APR-2001  (first entry)
XX
DE   Nucleotide sequence of a human SGT4-2 polypeptide.
XX
KW   Human; SGT4; signal transduction; guanosine triphosphate binding protein;
KW   GTP binding protein; cancer; immune response; nutritional source;
KW   animal feed; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..681
FT                   /*tag=  a
FT                   /product= "SGT4"
XX
PN   WO200078959-A1.
XX
PD   28-DEC-2000.
XX
PF   22-JUN-2000; 2000WO-US17248.
XX
PR   23-JUN-1999;   99US-0140627.
XX
PA   (LEXI-) LEXICON GENETICS INC.
XX
PI   Turner AC,  Zambrowicz B,  Nehls M,  Friedrich GA,  Sands AT;
XX
DR   WPI; 2001-032329/04.
DR   P-PSDB; AAB31564.
XX
PT   New SGT4 genes and proteins, useful for diagnosing and treating
```

PT disorders involving inappropriate regulation of a signal transduction
PT mechanism e.g. cancer -
XX
PS Claim 1; Fig 3; 82pp; English.
XX
CC The present sequence encodes a human SGT4 polypeptide. SGT4 polypeptides
CC are involved in signal transduction pathways regulated by guanosine
CC triphosphate (GTP) binding proteins). SGT4 polynucleotides and
CC polypeptides are for diagnosing and treating conditions related to a
CC signal transduction mechanism involving SGT4 such as cancer. In
CC addition, it can be used to detect the expression of SGT4 as markers of
CC specific cells and tissues such as neuronal tissue, heart, liver,
CC pancreas and adrenal gland. They are also useful for the construction of
CC transgenic and knockout animals for studying SGT4 function in vivo and
CC for the screening of SGT4 (ant)agonists in an animal model. Other more
CC general uses include: as molecular weight markers on Southern gels; as
CC chromosome markers or tags; as probes; for selecting and making
CC oligomers for attachment to a gene chip; to raise anti-protein or
CC anti-DNA antibodies or to elicit immune response. They are also
CC also be used as nutritional sources or supplements such as in animal
CC feed.
XX
SQ Sequence 681 BP; 212 A; 138 C; 142 G; 189 T; 0 other;

Query Match 61.0%; Score 681; DB 22; Length 681;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    436 atgagaattctggatctgccaaaaaaccaaatctcacatcttccagcagaaatcggttgt 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgagaattctggatctgccaaaaaaccaaatctcacatcttccagcagaaatcggttgt 60

Qy    496 ttgaagaacctgaaagaactcaatgtgggtttcaactatctgaagagcattcctccagaa 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ttgaagaacctgaaagaactcaatgtgggtttcaactatctgaagagcattcctccagaa 120

Qy    556 ttgggagattgtgaaaatctagagagactggattgttctggaaatctagaattaatggag 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ttgggagattgtgaaaatctagagagactggattgttctggaaatctagaattaatggag 180

Qy    616 ctgccctttgaattaagtaatttgaagcaagttacatttgtagatatctcagcaaacaag 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ctgccctttgaattaagtaatttgaagcaagttacatttgtagatatctcagcaaacaag 240

Qy    676 ttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcagtggttggatatcagc 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ttttccagtgtcccaatctgtgtcctgcggatgtcgaatttgcagtggttggatatcagc 300

Qy    736 agcaataacctgaccgacctgccgcaagatatagacaggctagaggagctgcagagcttt 795
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 agcaataacctgaccgacctgccgcaagatatagacaggctagaggagctgcagagcttt 360

Qy    796 ctcttgtataaaaacaagttgacctaccttccctattccatgctgaacctgaagaagctc 855
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ctcttgtataaaaacaagttgacctaccttccctattccatgctgaacctgaagaagctc 420

Qy    856 actctgttagtcgtcagtggggaccatttggtggagctcccaactgccctttgtgactca 915
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 actctgttagtcgtcagtggggaccatttggtggagctcccaactgccctttgtgactca 480

Qy    916 tccacacctttaaaatttgtaagccttatggacaatcctattgataatgcccaatgtgaa 975
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 tccacacctttaaaatttgtaagccttatggacaatcctattgataatgcccaatgtgaa 540

Qy    976 gatggcaatgaaataatggaaagtgaacgggatcgccaacattttgataaagaagttatg 1035
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 gatggcaatgaaataatggaaagtgaacgggatcgccaacattttgataaagaagttatg 600

Qy   1036 aaagcctatattgaagaccttaaagaaagagaatctgttcccagctataccaccaaagtg 1095
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 aaagcctatattgaagaccttaaagaaagagaatctgttcccagctataccaccaaagtg 660
```

```
Qy   1096 tcttttagccttcaactttga 1116
          ||||||||||||||||||||
Db    661 tcttttagccttcaactttga 681
```

RESULT 3
AAH17218
ID AAH17218 standard; cDNA; 2056 BP.
XX
AC AAH17218;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human cDNA sequence SEQ ID NO:16594.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 16594; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2056 BP; 642 A; 394 C; 495 G; 525 T; 0 other;

Query Match 11.6%; Score 129; DB 22; Length 2056;
Best Local Similarity 100.0%; Pred. No. 2.2e-55;

```
  Matches  129;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     988 ataatggaaagtgaacgggatcgccaacattttgataaagaagttatgaaagcctatatt 1047
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 ataatggaaagtgaacgggatcgccaacattttgataaagaagttatgaaagcctatatt 1353

Qy    1048 gaagaccttaaagaaagagaatctgttcccagctataccaccaaagtgtcttttagcctt 1107
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 gaagaccttaaagaaagagaatctgttcccagctataccaccaaagtgtcttttagcctt 1413

Qy    1108 caactttga 1116
           |||||||||
Db    1414 caactttga 1422


RESULT   4
AAT41852
ID   AAT41852 standard; DNA; 9789 BP.
XX
AC   AAT41852;
XX
DT   20-FEB-1997  (first entry)
XX
DE   cDNA encoding Plasmodium falciparum erythrocyte membrane protein.
XX
KW   Plasmodium falciparum; erythrocyte membrane protein; malaria;
KW   detection; identification; treatment; prevention; parasite; ss.
XX
OS   Plasmodium falciparum MC type.
XX
FH   Key             Location/Qualifiers
FT   CDS             326..9497
FT                   /*tag= a
FT                   /product= Erythrocyte membrane protein
FT   misc_feature    518..520
FT                   /*tag= b
FT                   /transl_except= GTA encodes Tyrosine
FT   misc_feature    656..658
FT                   /*tag= c
FT                   /transl_except= ATT encodes Leucine
FT   misc_feature    2909..2911
FT                   /*tag= d
FT                   /transl_except= AAC encodes Aspartic acid
FT   misc_feature    3461..3463
FT                   /*tag= e
FT                   /transl_except= GAA encodes Glutamine
FT   misc_feature    5546..5548
FT                   /*tag= f
FT                   /transl_except= CCT encodes Arginine
FT   misc_feature    6254..6256
FT                   /*tag= g
FT                   /transl_except= AAT encodes Lysine
FT   misc_feature    6257..6259
FT                   /*tag= h
FT                   /transl_except= ATA encodes Tyrosine
FT   misc_feature    6263..6265
FT                   /*tag= i
FT                   /transl_except= AAC encodes Lysine
FT   misc_feature    6269..6271
FT                   /*tag= j
FT                   /transl_except= TTC encodes Isoleucine
FT   misc_feature    6272..6274
FT                   /*tag= k
FT                   /transl_except= ATA encodes Histidine
FT   misc_feature    6275..6277
FT                   /*tag= l
FT                   /transl_except= ATT encodes Asparagine
FT   misc_feature    6278..6280
FT                   /*tag= m
FT                   /transl_except= GGA encodes Tryptophan
FT   intron          7754..8478
FT                   /*tag= n
XX
PN   WO9633736-A1.
XX
PD   31-OCT-1996.
XX
PF   26-APR-1996;   96WO-US05798.
XX
PR   27-APR-1995;   95US-0430908.
XX
PA   (AFFY-) AFFYMAX TECHNOLOGIES NV.
XX
PI   Baruch DI,  Howard RJ,  Pasloske BL;
XX
DR   WPI; 1996-497376/49.
DR   P-PSDB; AAW00384.
XX
PT   New Plasmodium falciparum erythrocyte membrane proteins - used to
PT   develop products for the diagnosis, treatment or prevention of
PT   malaria parasite infections
XX
PS   Disclosure; Figure 12; 149pp; English.
XX
CC   A polypeptide comprising a Plasmodium falciparum (Pf) erythrocyte
CC   membrane protein 1 (PfEMP1) or active fragments or analogues of that
CC   protein can be used in the treatment or prevention of symptoms of a
CC   malaria parasite infection.  The polypeptides can inhibit, block or
CC   reverse the sequestration of erythrocytes in patients suffering from
CC   malaria. Nucleic acids derived from the PfEMP1 gene can be used as
CC   probes and primers to identify a Plasmodium falciparum parasite, the
CC   primers used to generate characteristic amplification patterns from
CC   different P. falciparum strains. Antibodies specifically
CC   immunoreactive with the PfEMP1 polypeptide or its fragments may be
CC   used in diagnosis of malaria infection. This sequence encodes the
CC   PfEMP1 protein of the MC type of Plasmodium falciparum.  An
CC   alternative, truncated version of the coding sequence (a cDNA clone)
CC   is given in AAT41853.
XX
SQ   Sequence 9789 BP; 4061 A; 1393 C; 1837 G; 2498 T; 0 other;


  Query Match             2.1%;  Score 23;  DB 17;  Length 9789;
  Best Local Similarity   100.0%;  Pred. No. 0.13;
  Matches   23;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     118 ttggagaagataaaggaggagtg 140
           |||||||||||||||||||||||
Db    3902 ttggagaagataaaggaggagtg 3924


RESULT   5
AAC46422/c
ID   AAC46422 standard; DNA; 1421 BP.
XX
AC   AAC46422;
XX
DT   18-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 50080.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
```

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 20; DB 21; Length 1421;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy     410 ctacatatattcagttattt 429
           ||||||||||||||||||||
Db    1343 CTACATATATTCAGTTATTT 1324
```

RESULT 6
AAC36467/c
ID AAC36467 standard; DNA; 1422 BP.
XX
AC AAC36467;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 13915.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 20; DB 21; Length 1422;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 410 ctacatatattcagttattt 429
||||||||||||||||||||
Db 1342 CTACATATATTCAGTTATTT 1323

RESULT 7
AAC00689
ID AAC00689 standard; cDNA; 327 BP.
XX
AC AAC00689;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 687.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
DR P-PSDB; AAG00683.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 687; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs

```
CC   derived from 30 different tissues. EST sequences usually correspond
CC   mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC   often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC   well suited for isolating cDNA sequences derived from the 5' ends of
CC   mRNAs and even in those cases where longer cDNA sequences have been
CC   obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC   mRNAs with intact 5' ends and can therefore be used to obtain full length
CC   cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC   gene therapy and chromosome mapping procedures. They are used to obtain
CC   upstream regulatory sequences and to design expression and secretion
CC   vectors.
XX
SQ   Sequence 327 BP; 121 A; 48 C; 85 G; 73 T; 0 other;


  Query Match             1.7%;  Score 19;  DB 21;  Length 327;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    497 tgaagaacctgaaagaact 515
          ||||||||||||||||||
Db     88 tgaagaacctgaaagaact 106


RESULT   8
AAC06478
ID   AAC06478 standard; cDNA; 413 BP.
XX
AC   AAC06478;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 10553.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 10553; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 413 BP; 138 A; 76 C; 103 G; 95 T; 1 other;


  Query Match             1.7%;  Score 19;  DB 21;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    497 tgaagaacctgaaagaact 515
          ||||||||||||||||||
Db    173 tgaagaacctgaaagaact 191


RESULT   9
AAT03478
ID   AAT03478 standard; DNA; 1185 BP.
XX
AC   AAT03478;
XX
DT   06-JUN-1996  (first entry)
XX
DE   Transcription factor BTF2 complex p44 subunit gene.
XX
KW   Transcription factor; BTF2; subunit; kinase; ATPase; helicase; HeLa; PCR;
KW   reconstitution; in vitro transcription system; probe; primer; antibody;
KW   amplification; microsequence; cancer; skin melanoma; xeroderma; UV light;
KW   Cockayne syndrome; skin pigmentation disorder; sensitivity; ss.
XX
OS   Homo sapiens.
XX
PN   WO9529245-A2.
XX
PD   02-NOV-1995.
XX
PF   25-APR-1995;   95WO-FR00540.
XX
PR   25-APR-1994;   94FR-0004937.
XX
PA   (ASRE-) ASSOC DEV RECH & GENETIQUE MOLECULAIRE.
XX
PI   Egly J,  Humbert S,  Moncollin V;
XX
DR   WPI; 1995-382993/49.
DR   P-PSDB; AAR88225.
XX
PT   New protein sub-unit(s) of transcription factor BTF2 - useful for
PT   treating or diagnosing deficiencies in DNA repair processes
XX
PS   Claim 1; Fig 2; 16pp; French.
XX
CC   This is the nucleotide sequence of the transcription factor BTF2 p44
CC   subunit gene.  The sequence encodes a protein of 395 amino acids.
CC   The genes for the p34 (AAT03477) and p44 subunits were isolated from a
CC   HeLa DNA library in lambda-ZAPII using oligonucleotide probes and
CC   primers based on microsequencing of the purified subunits (e.g.
CC   AAT03479-80). Neither the p34 nor the p44 subunits contain any kinase,
CC   ATPase or helicase activity and cannot be used to reconstitute BTF2
CC   activity even with the p62 and p89 BTF2 subunits in an in vitro
CC   transcription system. The proteins can be used to raise antibodies useful
CC   for detecting abnormally low levels of the subunits.  The DNA sequences
CC   can be used similarly for DNA levels. The antibodies and probes are
CC   useful in the detection of development of cancer, partic. skin melanoma
CC   but also xeroderma or Cockayne syndrome, skin pigmentation disorders or
CC   sensitivity to UV light.
XX
SQ   Sequence 1185 BP; 356 A; 220 C; 255 G; 354 T; 0 other;


  Query Match             1.7%;  Score 19;  DB 16;  Length 1185;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    497 tgaagaacctgaaagaact 515
```

```
          ||||||||||||||||||
Db      6 tgaagaacctgaaagaact 24


RESULT  10
AAC78156
ID   AAC78156 standard; cDNA; 1816 BP.
XX
AC   AAC78156;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human cancer associated gene sequence SEQ ID NO:550.
XX
KW   Human; cancer associated gene; cancer antigen; detection; cancer;
KW   diagnosis; cytostatic; proliferative; vulnerary; immunomodulator;
KW   antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
KW   antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
KW   dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW   vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW   immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW   allergic reaction; graft versus host disease; organ rejection;
KW   haemostatic; thrombolytic; cardiovascular disorder; infection;
KW   neurological disease; drug screening; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055350-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05882.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-587533/55.
DR   P-PSDB; AAB43947.
XX
PT   Novel isolated nucleic acids comprising sequences encoding peptides
PT   useful for treating or diagnosing e.g. cancer -
XX
PS   Claim 1; Page 1075-1076; 2352pp; English.
XX
CC   AAC77607 to AAC78448 encode the human cancer associated proteins given
CC   in AAB43398 to AAB44239. The proteins can have activities based on the
CC   tissues and cells the genes are expressed in. Example of activities
CC   include: cytostatic; proliferative; vulnerary; immunomodulator;
CC   antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC   antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC   dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC   nootropic; vasotropic; antipsoriatic and antiangiogenic. The
CC   polynucleotides and polypeptides can be used for preventing, treating or
CC   ameliorating medical conditions and diagnosing pathological conditions.
CC   Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC   the present invention may be used to treat immune disorders by activating
CC   or inhibiting the proliferation, differentiation or mobilisation of
CC   immune cells, to treat disorders of haematopoietic cells, autoimmune
CC   disorders, allergic reactions, graft versus host disease and organ
CC   rejection, modulate haemostatic or thrombolytic activity, modulate
CC   inflammation, cancers, cardiovascular disorders, neurological disease and
CC   bacterial or viral infections. The peptides, nucleotides, antibodies,
CC   agonists and antagonists may be also be used in drug screens. AAC78449 to
CC   AAC78457 and AAB44240 represent sequences used in the exemplification of
CC   the present invention.
XX
SQ   Sequence 1816 BP; 561 A; 292 C; 398 G; 563 T; 2 other;

  Query Match             1.7%;  Score 19;  DB 21;  Length 1816;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     497 tgaagaacctgaaagaact 515
           ||||||||||||||||||
Db     223 tgaagaacctgaaagaact 241


RESULT  11
AAA75580
ID   AAA75580 standard; DNA; 3072 BP.
XX
AC   AAA75580;
XX
DT   22-JAN-2001  (first entry)
XX
DE   DNA encoding a mouse zalpha11 ligand polypeptide.
XX
KW   zalpha11 ligand; cytokine; haematopoietic cell proliferation; lymphoma;
KW   tumourigenesis; leukaemia; hematopoiesis; B cell tumour; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             54..494
FT                   /*tag=  a
FT                   /product= "Zalpha11"
XX
PN   WO200053761-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06067.
XX
PR   09-MAR-1999;   99US-0264908.
PR   11-MAR-1999;   99US-0265992.
PR   01-JUL-1999;   99US-0142013.
XX
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Novak JE,  Presnell SR,  Sprecher CA,  Foster DC,  Holly RD,  Gross JA;
PI   Johnston JV,  Nelson AJ,  Dillon SR,  Hammond AK;
XX
DR   WPI; 2000-565600/52.
DR   P-PSDB; AAB18624.
XX
PT   New human cytokine, designated zalpha11 ligand, useful for stimulating
PT   the proliferation and/or development of haematopoietic cells in vitro
PT   and in vivo, and for treating tumourigenesis -
XX
PS   Disclosure; Page 220-222; 256pp; English.
XX
CC   The present sequence encodes a mouse zalpha11 ligand polypeptide,
CC   which is a cytokine. The zalpha11 ligand is useful for stimulating the
CC   proliferation and development of haematopoietic cells in vitro and in
CC   vivo. Zalpha11 ligand polynucleotides can be used as primers or probes
CC   for cloning the zalpha11 gene. The zalpha11 ligand is useful for
CC   treating tumourigenesis. A zalpha11 ligand-saporin fusion toxin may be
CC   used for treating leukaemias and lymphomas. Antagonists against zalpha11
CC   ligand are useful as research reagents for characterizing ligand-receptor
CC   interaction. Antagonists are also useful for inhibiting expansion,
CC   proliferation, activation and differentiation of cells involved in
CC   regulating hematopoiesis. The zalpha11 ligand may also be used to
CC   stimulate an immune response against B cell tumour, a virus, a parasite
CC   or a bacterium. The zalpha11 polypeptides, polynucleotides, antagonists,
CC   agonists and antibodies are also useful for the detection, diagnosis,
CC   prevention, and treatment of diseases associated with a zalpha11 ligand
CC   genetic defect.
XX
SQ   Sequence 3072 BP; 925 A; 591 C; 623 G; 933 T; 0 other;

  Query Match             1.7%;  Score 19;  DB 21;  Length 3072;
```

```
  Best Local Similarity   100.0%;  Pred. No. 14;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     970 tgtgaagatggcaatgaaa 988
           ||||||||||||||||||||
Db    1257 tgtgaagatggcaatgaaa 1275


RESULT  12
AAX13947
ID   AAX13947 standard; DNA; 811 BP.
XX
AC   AAX13947;
XX
DT   31-MAR-1999  (first entry)
XX
DE   H. pylori GHPO 1275 gene.
XX
KW   GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW   peptic ulcer disease; ss.
XX
OS   Helicobacter pylori.
XX
FH   Key             Location/Qualifiers
FT   CDS             51..764
FT                   /*tag= a
XX
PN   WO9843478-A1.
XX
PD   08-OCT-1998.
XX
PF   01-APR-1998;   98WO-US06371.
XX
PR   29-JUL-1997;   97US-0902615.
PR   01-APR-1997;   97US-0833457.
PR   24-JUN-1997;   97US-0881227.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (INMR ) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
PI   Al-Garawi A,  Kleanthous H,  Miller C,  Oomen RP,  Tomb J;
XX
DR   WPI; 1998-542293/46.
DR   P-PSDB; AAW98228.
XX
PT   New isolated Helicobacter polynucleotides - used to develop products
PT   for the diagnosis, prevention and treatment of Helicobacter
PT   infections and gastrointestinal diseases
XX
PS   Claim 1; Page 164-165; 2054pp; English.
XX
CC   This sequence represents a polynucleotide of the invention. It was
CC   isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.
CC   The polypeptides can be used for preventing or treating Helicobacter
CC   infections, and gastroduodenal diseases associated with these
CC   infections, including acute, chronic, and atrophic gastritis, and peptic
CC   ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC   for the production of antibodies. The products can also be used for
CC   detection and diagnosis.
XX
SQ   Sequence 811 BP; 253 A; 146 C; 187 G; 225 T; 0 other;


  Query Match             1.6%;  Score 18;  DB 19;  Length 811;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     915 atccacacctttaaaatt 932
           ||||||||||||||||||
Db     729 atccacacctttaaaatt 746


RESULT  13
```

```
AAT66241
ID   AAT66241 standard; cDNA; 1981 BP.
XX
AC   AAT66241;
XX
DT   28-JUL-1997  (first entry)
XX
DE   Romaine lettuce violaxanthin de-epoxidase cDNA.
XX
KW   Violaxanthin de-epoxidase; VDE; light; photosensitivity;
KW   photoprotection; transgenic plant; zeaxanthin; antheraxanthin;
KW   xanthophyll; lettuce; ss.
XX
OS   Lactuca sativa L. cv. romaine.
XX
FH   Key             Location/Qualifiers
FT   misc_difference 26..29
FT                   /*tag=  a
FT                   /note= "bases 26-29 are illegible in Fig 1"
FT   misc_difference 66..72
FT                   /*tag=  b
FT                   /note= "bases 66-72 are illegible in Fig 1"
FT   misc_difference 105..110
FT                   /*tag=  c
FT                   /note= "bases 105-110 are illegible in Fig 1"
FT   misc_difference 147..149
FT                   /*tag=  d
FT                   /note= "bases 147-149 are illegible in Fig 1"
FT   misc_difference 186..189
FT                   /*tag=  e
FT                   /note= "bases 186-189 are illegible in Fig 1"
FT   misc_difference 226..227
FT                   /*tag=  f
FT                   /note= "bases 226-227 are illegible in Fig 1"
FT   CDS             235..1656
FT                   /*tag=  g
FT   transit_peptide 235..609
FT                   /*tag=  h
FT   mat_peptide     610..1653
FT                   /*tag=  i
XX
PN   WO9717447-A2.
XX
PD   15-MAY-1997.
XX
PF   07-NOV-1996;   96WO-US18291.
XX
PR   06-AUG-1996;   96US-0023502.
PR   07-NOV-1995;   95US-0006315.
XX
PA   (CALJ ) CALGENE INC.
XX
PI   Bugos RC,  Rockholm DC,  Yamamoto HY;
XX
DR   WPI; 1997-281036/25.
DR   P-PSDB; AAW09874.
XX
PT   DNA encoding plant violaxanthin de-epoxidase - used to modify the
PT   sensitivity of a plant to light
XX
PS   Claim 3; Fig 1; 41pp; English.
XX
CC   A cDNA clone (AAT66241) codes for the 55 kDa violaxanthin de-epoxidase
CC   (VDE) (AAW09874) of romaine lettuce.  VDE was purified from romaine
CC   lettuce chloroplasts and 2 tryptic peptides were used to develop
CC   primers (see also AAT66244-45), which amplified a partial VDE
CC   sequence.  The amplified sequence was then used to screen a lettuce
CC   cDNA library, and the 1981 bp DNA sequence was identified.  VDE
CC   nucleic acids (see also AAT66242-43), in sense or antisense
CC   orientation, can be used in genetic constructs, pref. also contg. a
CC   plastid translocation sequence, to modify VDE levels in plants.
CC   Increased levels result in the plant being tolerant of increased
CC   light and therefore more productive and/or more resistant to
```

```
CC   disease.  Underexpression of VDE increases photosynthetic
CC   efficiency under low light.  The photosensitivity of a range of
CC   crops, trees and ornamentals can be modified.
XX
SQ   Sequence 1981 BP; 608 A; 337 C; 433 G; 577 T; 26 other;


  Query Match             1.6%;  Score 18;  DB 18;  Length 1981;
  Best Local Similarity   100.0%;  Pred. No. 43;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      93 ggtggaaaggcttgagaa 110
           ||||||||||||||||||
Db    1371 ggtggaaaggcttgagaa 1388


RESULT  14
AAD10125/c
ID   AAD10125 standard; cDNA; 2324 BP.
XX
AC   AAD10125;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Mouse serotransferrin (siderophilin) cDNA.
XX
KW   Mouse; cytostatic; antiinflammatory; immunoregulatory; tissue integrity;
KW   wound healing; immune response; vaccine; cancer; asthma; allergy;
KW   cell trafficking; therapy; secreted protein; serotransferrin;
KW   siderophilin; Tf; beta-1-metal binding globulin; transferrin; ss.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             43..2136
FT                   /*tag= a
FT                   /product= "Mouse serotransferrin (siderophilin)"
XX
PN   WO200148192-A1.
XX
PD   05-JUL-2001.
XX
PF   21-DEC-2000; 2000WO-NZ00256.
XX
PR   23-DEC-1999;   99US-0171678.
PR   28-NOV-2000; 2000US-0724864.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI   Watson JD,  Murison JG;
XX
DR   WPI; 2001-425665/45.
DR   P-PSDB; AAE05358.
XX
PT   Novel isolated polypeptide useful to isolate corresponding interacting
PT   proteins or other compounds, to quantitatively determine levels of
PT   interacting proteins or other compounds, and as therapeutic target -
XX
PS   Claim 1; Page 61-62; 101pp; English.
XX
CC   The patent discloses novel polynucleotides and their corresponding
CC   proteins which play a major role in induction of growth, cell migration
CC   and proliferation, cell-cell interaction and the differentiation of
CC   tissue-specific cells. These proteins are important in the maintenance
CC   of tissue integrity and thus are important in wound healing. They are
CC   useful in various assays to determine the biological activity, to raise
CC   antibodies, to isolate corresponding interacting proteins or other
CC   compounds, to quantitatively determine levels of interacting proteins or
CC   other compounds, and as therapeutic target in a whole range of disease
CC   states. Compositions comprising the novel proteins of the invention are
CC   useful for treating mammalian disorders. Polynucleotides of the invention
CC   are useful in genome and physical mapping, in positional cloning of
CC   genes, to tag or identify an organism or its reproductive material (as
CC   non-disruptive tags for marking organisms), and for the diagnosis and
CC   treatment of mammalian diseases which is the consequence of inappropriate
CC   expression of kinase genes. They are useful for promoting immune response
CC   as part of a vaccine or anti-cancer treatment, as target for cancer
CC   treatment, as immunoregulatory and anti-inflammatory molecule, as
CC   diagnostic for specific types of cancer and for development of an
CC   anti-cancer treatment, and as a target for antagonists in the treatment
CC   of diseases such as asthma and allergy. They are also useful to inhibit
CC   or enhance the activity of the soluble molecule that binds proteins of
CC   the invention, for tissue and neural regeneration, to promote or block
CC   cell trafficking, and as anti-inflammatory and/or vaccine adjuvant.
CC   The present sequence is a cDNA encoding mouse serotransferrin
CC   (siderophilin). Serotransferrin (Tf) also known as beta-1-metal
CC   binding globulin is a part of the transferrin family.
XX
SQ   Sequence 2324 BP; 592 A; 585 C; 627 G; 520 T; 0 other;


  Query Match             1.6%;  Score 18;  DB 22;  Length 2324;
  Best Local Similarity   100.0%;  Pred. No. 44;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     709 tcgaatttgcagtggttg 726
           ||||||||||||||||||
Db    1538 TCGAATTTGCAGTGGTTG 1521


RESULT  15
AAV26082/c
ID   AAV26082 standard; cDNA; 3997 BP.
XX
AC   AAV26082;
XX
DT   07-JUN-1999  (first entry)
XX
DE   Tomato pest resistance Mi gene (copy 1).
XX
KW   Pest resistance; nematode resistance; disease resistance; Mi gene;
KW   tomato; transgenic plant; crop protection; biological control; ss.
XX
OS   Lycopersicon esculentum.
XX
FH   Key             Location/Qualifiers
FT   CDS             85..3852
FT                   /*tag= a
XX
PN   WO9815171-A1.
XX
PD   16-APR-1998.
XX
PF   09-OCT-1997;   97WO-US18802.
XX
PR   10-OCT-1996;   96US-0028191.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Bodeau J,  Kaloshian I,  Milligan S,  Williamson VM;
PI   Yaghoobi J;
XX
DR   WPI; 1998-240529/21.
DR   P-PSDB; AAW55974.
XX
PT   Nucleic acids encoding Mi polypeptide(s) conferring nematode
PT   resistance - useful to produce transgenic plants resistant to these
PT   and other pests, and in marker-aided selection to assess cultivars
PT   for resistance
XX
PS   Claim 1; Page 40-42; 55pp; English.
XX
CC   This is the nucleotide sequence of a cDNA clone encoding tomato Mi
CC   polypeptide (see AAW55974), which confers resistance to nematodes
CC   such as Globodera, Heterodera and Meloidogyne spp., and other pests
CC   such as aphids.  A genetic locus, Mi, was localised by genetic
```

CC analysis to a region of the tomato genome of about 65 kb. DNA
CC corresponding to this region was cloned into bacterial artificial
CC chromosome vectors. Sequence analysis of a 52 kb BAC3 insert (see
CC AAV26084) identified 3 open reading frames, one of which is probably
CC a pseudogene. By RNA blot analysis, transcripts of approximately
CC 4 kb corresponding to copy 1 and copy 2 were found in both
CC resistant and susceptible tomato roots and in leaves of resistant but
CC not susceptible plants. cDNA sequences corresponding to full-length
CC transcripts of copy 1 (i.e. AAV26082) and copy 2 (see AAV26083) were
CC obtained. The encoded polypeptides are 91% identical and contain
CC structural features similar to known plant resistance genes (R
CC genes) of the nucleotide binding site/leucine-rich repeat (NBS/LRR)
CC family. A recombinant expression cassette comprising an Mi
CC polynucleotide and an operably linked plant promoter can be used
CC to enhance nematode resistance in plants especially tomatoes
CC (claimed). Transgenic plants can also be constructed using a Mi
CC promoter with heterologous genes. Mi polynucleotides are useful in
CC marker-aided selection, providing a means of DNA fingerprinting
CC cultivars and wild germplasm with respect to their disease
CC resistance.
XX
SQ Sequence 3997 BP; 1252 A; 664 C; 852 G; 1228 T; 1 other;

Query Match 1.6%; Score 18; DB 19; Length 3997;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 299 cgtttgtgtttgaacttt 316
||||||||||||||||||
Db 3370 CGTTTGTGTTTGAACTTT 3353

Search completed: February 26, 2002, 13:21:28
Job time: 5604 sec